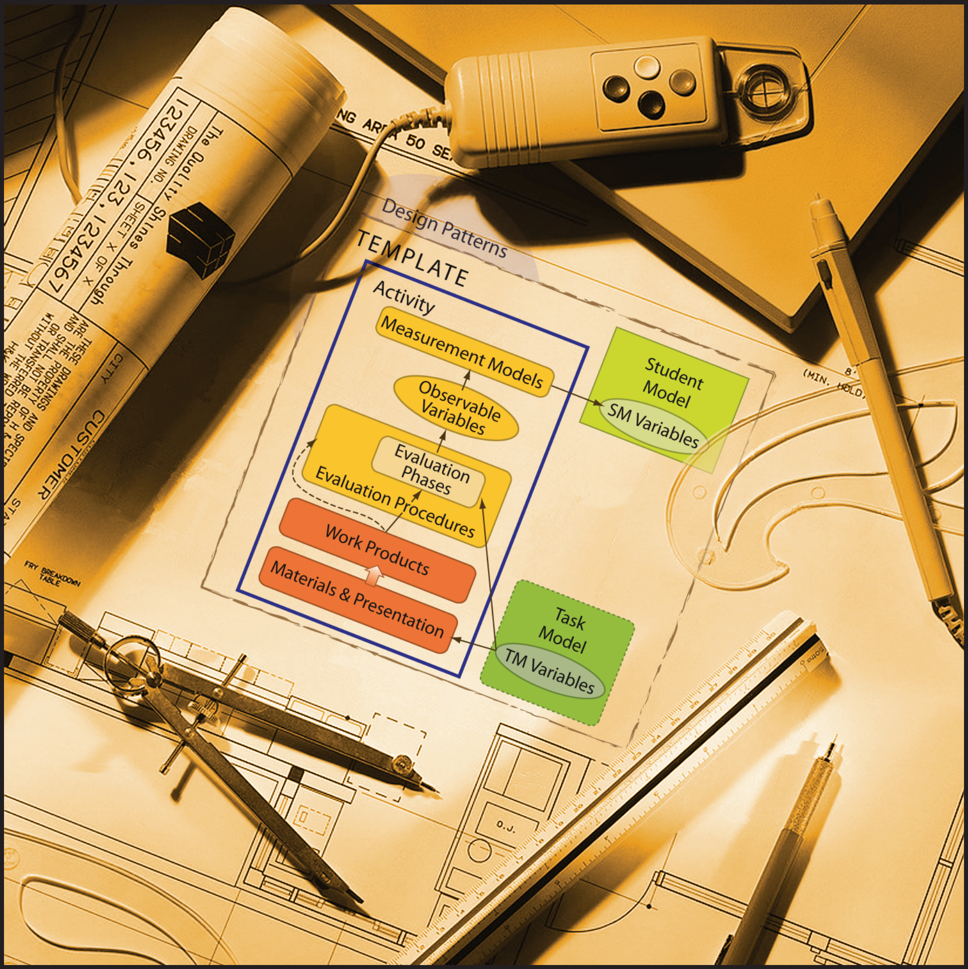


Large-Scale Assessment Technical Report 10|
November 2010

Using Evidence-Centered Design to Support Assessment, Design, and Validation of Learning Progressions

Project: Application of Evidence-Centered Design to State Large-Scale Science Assessment

Daniel Zalles and Geneva Haertel, SRI International
Robert J. Mislevy, University of Maryland

Report Series Published by SRI International

SRI International
Center for Technology in Learning
333 Ravenswood Avenue
Menlo Park, CA 94025-3493
650.859.2000
http://ecd.sri.com

EVIDENCE-CENTERED DESIGN TO STATE LARGE-SCALE SCIENCE ASSESSMENT
TECHNICAL REPORT 10

Using Evidence-Centered Design to Support Assessment, Design, and Validation of Learning Progressions

November 2010

Prepared by:
Daniel Zalles and Geneva Haertel, SRI International
Robert J. Mislevy, University of Maryland

Acknowledgments
This material is based on work supported by the National Science Foundation under grant DRL-0733172 (An Application of Evidence-Centered Design to State Large-Scale Science Assessment).

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This report explores how evidence-centered design (ECD) can be applied to the development of student tasks that can be used to assess the evolution of student reasoning in instructional programs designed around learning progressions. To illustrate the applicability of ECD, a *design pattern* was constructed that captured the *domain analysis* articulated in a published paper describing a learning progression about the study of modern genetics (Duncan, Rogat, and Yarden, 2009). Both the learning progression and the *design pattern* maintain three grade level bands (5-6, 7-8, and 9-10) and concern the acquisition of student understanding about the genetic, molecular, and meiotic models across the bands. The report also addresses related issues that bear on ECD's usefulness for learning progression student assessment using examples of assessment tasks designed by the first author. We discuss ways in which *design pattern* structures can be varied to reflect different ways in which the components and boundaries of learning progressions can be demarcated and how ECD can be used to examine the strength of alignments between the learning progressions and standards. Finally, the report identifies the challenges faced by researchers working to validate their learning progressions.

1.0 Introduction

A panel composed of twenty-one science educators, learning scientists, psychologists, assessment experts, policy researchers, and curriculum developers were convened by the Consortium for Policy Research and Education at Teachers College, Columbia University to "review and discuss the state of the current work on learning progressions" (Corcoran. Mosher, and Rogat, 2009, p. 5). They agreed to a definition of learning progressions as "hypothesized descriptions of the successively more sophisticated ways student thinking about how an important domain of knowledge or practice develops as children learn about and investigate that domain over an appropriate span of time" (Ibid., p. 37). In addition, they agreed that learning progressions are grounded in the hypothesis that "most students' understanding will move through these intermediate conceptions in roughly the same order, though perhaps at quite different rates (depending on instruction, ability, other experiences and exposure, including home opportunities etc.)" (Ibid., p. 42). Hence, learning progressions start at the most naive student conceptions and end at the most sophisticated and accurate understandings about focal scientific phenomena. In science domains, the focus of the learning progression may be student reasoning about a particular scientific phenomenon or principle (i.e., the "content"), or about what characterizes student capability to practice what characterizes scientific inquiry (such as model-based reasoning), or about both.

Hence, a learning progression represents the intersection between the epistemic characteristics of the focal reasoning and the cognitive developmental characteristics of the learner. The hypothesized utility of a learning progression is that instructional programs and their assessments will be more responsive to how students progress on their acquisition of a particular skill or understanding when the programs and assessments are explicitly designed around the learning progression. Learning progressions are currently being developed in a number of science content domains, such as atomic molecular theory (Smith et al., 2006), biodiversity (Songer et al., 2009), buoyancy (Kennedy and Wilson, 2007), carbon in ecosystems (Mohan, Chen, and Anderson, 2009), and movement of water through the environment (Gunckel, Covitt, and Anderson, 2009).

Learning progressions vary in how a particular science domain can be framed as teachable and assessable components or levels. Each learning progression, however, attempts to link forms of reasoning and cognition to the dimensions of what characterizes exceedingly deeper and more accurate understandings of the focal content, and, in some cases, skills that transcend specific scientific phenomena such as model-based reasoning (Schwarz et al., 2009). Some move beyond these attributes as well to incorporate pedagogical attributes. For example, a team of researchers who has been developing and testing a learning progression about biodiversity and evidence-based reasoning has embedded strategically situated scaffolds into the evidence-based reasoning track, designed to push students into progressively higher levels of evidence-based reasoning in relation to their levels of understanding about biodiversity (Songer et. al., 2009).

Learning progressions have the potential to transform educational practice by grounding instructional sequences in better understandings of the natural development of science learning. Assessments grounded in evidence-centered design (ECD; Mislevy, Steinberg, & Almond, 2003; Mislevy & Haertel, 2006) have the potential to capture the learning progressions' reasoning-oriented emphases in a targeted and valid way. The primary purpose of this report is to explore how ECD can be applied to the development of tasks that can be used to assess the evolution of student reasoning in instructional programs designed around learning progressions. This topic is addressed in Section 2.0. As models of evolving student reasoning about a particular content domain, learning progressions are amendable to being represented by the structures of ECD. Specifically, in science learning progressions, the characteristics of what is in ECD referred to as the "Student Model" are the students' evolving reasoning capacities about the focal scientific phenomena. We show how an ECD support tool called a *design pattern* (Mislevy, et al., 2003) can play a pivotal role in task design

Section 2.1 presents an example of how a *design pattern* can be constructed to reflect the dimensions of a particular learning progression in ways that can serve as a foundation for assessment item development. The example also illustrates how ECD principles can be applied to decision-making on appropriate scope and grain size for identifying the assessable constructs of the learning progression (e.g., the scope and span of the construct to be represented in a particular design pattern and the criteria for determining the extent to which levels in the learning progression can be separately yet interdependently modeled). Section 2.1 also presents examples of new assessment tasks that align to the different attributes of the *design pattern*.

The other subsections in Section 2.0 address related issues that bear on ECD's usefulness for learning progression student assessment. Section 2.2 presents examples of assessment tasks designed by the first author that respond to attributes of the Student Model expressed in the other fields of the *design pattern*. Section 2.3 discusses ways in which *design pattern* structures can be varied to reflect different ways in which the components and boundaries of learning progressions can be demarcated. Section 2.4 addresses how ECD can be used to examine the strength of alignments between the learning progressions and standards. These examinations can help determine to what extent standards-driven instructional programs and assessments would need to be redesigned to become more grounded in the learning progressions if teachers are ever going to be held accountable to teaching to the progressions.

Lastly, in Section 3.0, the report identifies the challenges faced by researchers working to validate their learning progressions. Validation methods used so far include cognitive interviews, longitudinal studies, and psychometric analyses of student responses to learning progression-grounded assessments. Section 2.0 will have argued that ECD can improve the validity of assessment tasks designed to measure impacts of learning progression-based instructional programs on student reasoning. Section 3.0 then argues that ECD also can facilitate greater alignment between assessment task's characteristics and learning progression constructs in ways that render the assessments more effective instruments for ascertaining the validity of the learning progressions themselves.

2.0 Applying ECD to Learning Progression Assessment Design

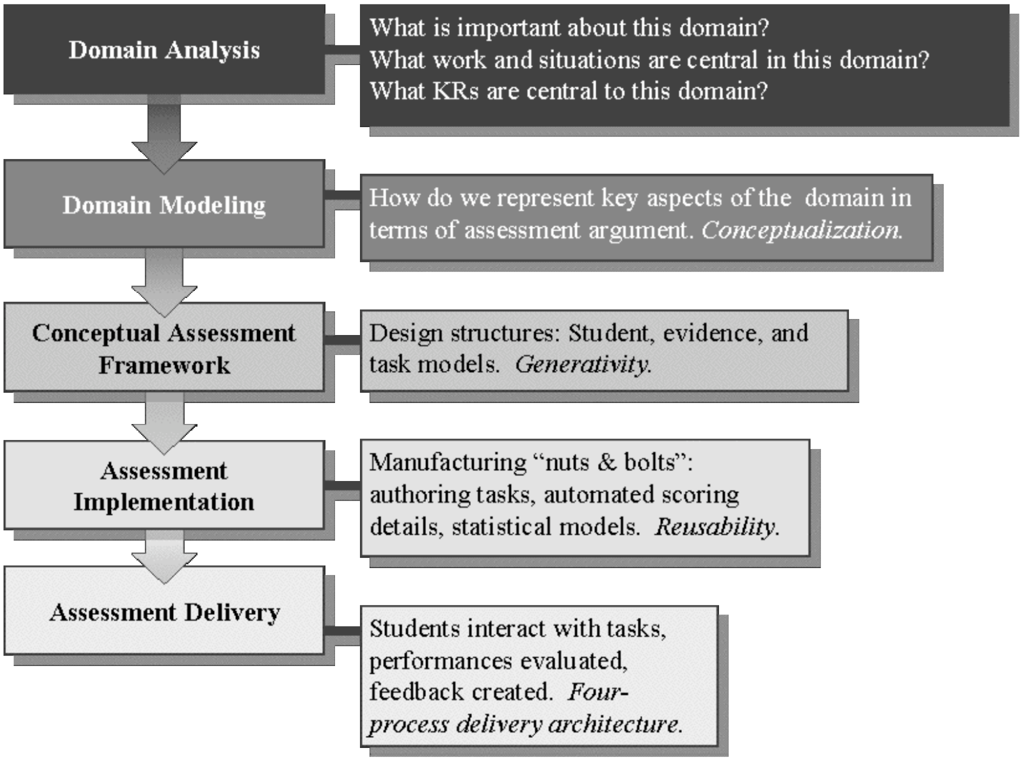
ECD is a framework for analyzing, designing, and implementing assessment from the perspective of assessment arguments. Messick succinctly captures the central idea (1994) in the following quote:

> A construct-centered approach would begin by asking what complex of knowledge, skills, or other attributes should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society. Next, what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors? Thus, the nature of the construct guides the selection or construction of relevant tasks as well as the rational development of construct-based scoring criteria and rubrics. (p. 17)

ECD lays out the activities, processes, and elements of the assessment enterprise in terms of the layers depicted in Figure 1. The first layer, *domain analysis*, concerns research and experience from the domain to be assessed, to examine what kinds of capabilities, situations, representations, interactions, and theoretical groundings will be needed to craft assessments around arguments. The research on learning progressions taking place in cognitive psychology and science education is a primary source of insight for assessment design from this perspective. The second layer, *domain modeling*, is about sketching out coherent assessment arguments in the domain, with more formal structuring of the elements in Messick's quotation. This is where the present report focuses its attention. The remaining layers involve more technical work in specifying blueprints for elements of measurement models, scoring procedures, task schemas, and the like; for implementing the elements; and for managing their interactions with one another and students in the actual operation of the assessment.

The *domain modeling* in ECD can catalyze valid measurement of student progress in instructional programs based around learning progressions. *Design patterns* provide structures that can articulate the levels of the learning progression and hence support the development of items that are grounded in the learning progression's Student Model.

Figure 1: Layers in Evidence-Centered Design



Applying ECD to the design of learning progression-grounded assessments can provide an alternative path for identifying assessable constructs that conform to the evolving reasoning foci of the learning progressions. The work of learning progression researchers reveals a shift in emphasis away from simply catalyzing and rewarding students' demonstrations of canonical knowledge and toward catalyzing and rewarding instead their capacities to express deep evolving reasoning. For example, in a paper that explores the instructional implications of teaching to their learning progression about the carbon cycle for grades 4-12, Mohan and Anderson (2009) suggest that instructional programs that do not provide opportunities for students to transition from everyday ("force-dynamic") discourse about the world to scientific discourse will fail to stimulate the students to reach the highest possible level in the learning progression—a level characterized by deep and authentic canonical understanding about the chemical properties and reactions that drive the carbon cycle. The implication is that in a scientific discourse-enabling classroom that aims to advance students from level I to level IV of this carbon cycle learning progression, the instructional program would need to be designed to respond to the developmental capabilities of the students to reason rather than to their ability to demonstrate surface-level knowledge. In Mohan and Anderson's terms, this "alternative pathway" to understanding the carbon cycle would transition students from "naming" the dimensions of the carbon cycle to "explaining" them. They write, "Naming focuses on key words and phrases characteristic of particular levels of reasoning. Explaining focuses on the structure of explanations, and how grounded these explanations are in terms of scientific principles" (Mohan and Anderson, p. 10). Duncan, Rogat, and Yarden (2009) adopt a similar perspective in their advocacy of instructional programs that look at the big picture about modern genetics rather than the typical emphasis on disconnected details that is offered in biology textbooks: an emphasis that was revealed in a 2006 AAAS textbook analysis study (AAAS, 2006). The dependency that growth of conceptual understanding has on growth of student capabilities to practice scientific discourse is also at the core of a learning progression about biodiversity (Songer et. al., 2009).

These learning progression researchers' perspectives suggest the need for two paradigm shifts in science instructional practice that would be compatible with ECD-derived assessments: (1) a shift away from judging students on their demonstrated ability to express canonical knowledge through naming without understanding, and toward judging them instead on their abilities to explain their thinking in a scientifically reasonable manner; and 2) a shift away from stress on content details to stress on big ideas. ECD can be helpful through the constructing of *design patterns* that are aligned to the canonical scientific understandings or inquiry skills expressed in standards yet inclusive of the levels of evolving reasoning that eventually should culminate in deep acquisition of the understandings and skills. ECD also can be applied to identifying constructs for measuring progress at different milestones along the progression, ranging from naïve forms of reasoning at one end to fully canonical skills and understandings on the other. The application of ECD to formulate milestone-sensitive items can help to ensure that the items reflect the underlying model of student reasoning at the different levels of the learning progression.

2.1 Learning Progression Design Pattern Example

Table 1 displays a *design pattern* rooted in elements of a particular learning progression about modern genetics.[1] The learning progression contains a relatively well-developed exposition of progressively more sophisticated understandings about genetics. It starts with the recognition that there exist genes and traits that organisms possess and inherit, and evolving into understanding that there can be mutations and that the genes have biological structures and functions at different levels of organization, from amino acids and proteins to body parts and whole organisms. The *design pattern* contains content from a journal article about the progression (Duncan, Rogat, & Yarden, 2009). The content has been placed into *design pattern* fields through a process of interpretation. In these fields, there are many direct quotes from the article, all referenced by the number of the page on which they appear.

---

[1] An on-line, interactive version of this design pattern can be found at http://design-drk.padi.sri.com/padi/do/AddNodeAction?NODE_ID=2233&state=viewNode

Both the learning progression and the *design pattern* maintain three grade level bands: grades 5-6, 7-8, and 9-10. Inferences were made by the authors about what to put in some of the *design pattern* fields that were not addressed in the publication. Texts that are the result of these inferences appear in Table 1 in italics.

The following paragraphs walk through the fields of the *design pattern*, describing their contents and their relationship to the assessment argument. To anticipate, the fields that specifically ground inferences about students' levels in learning progressions are: Focal Knowledge, Skills and Abilities (KSAs) describe the levels of the progression. Potential Observations describe the kinds of performances that students at different levels of the learning progression are likely to exhibit, in tasks with Characteristic Features selected to reveal the corresponding levels of thinking.

The *design pattern* contains a series of *construct relevant features* that respond to the Student Model at the different levels of the learning progression. Focal KSAs in a *design pattern* are the knowledge, skills, or other attributes that tasks are meant to provide evidence about, as the Messick quotation begins with—that is, the construct that is the target of assessment. In this example, the Focal KSAs are derived from passages in the article that identify how students come to understand in ever-more sophisticated ways the characteristics of genes, proteins, and the outcomes of genetic changes on cells and organisms. These particular components of the larger learning progression were selected because they are the most fully followed up in the article by information that can be used to build assessment arguments. Although the entire learning progression has eight main ideas, the Focal KSAs of this *design pattern* focus primarily on the 2nd and 3rd ones. For background however, all of the main ideas are presented in the Overview section details

The Characteristic Features express common foundational attributes of assessment tasks that would be appropriate for measuring student progress in the Focal KSAs at the three grade level bands. They provide information that relates to Messick's question about the situations that can elicit behaviors that tell us something about the student's thinking. Variable Features present aspects of tasks that a designer can vary further to focus attention on certain aspects of capabilities, make tasks harder or easier, or bring in or avoid other knowledge.

Two fields in the *design pattern* address Messick's question about what behaviors or performances we need to see as evidence of the student's thinking: Potential Observations and Potential Work Products. The Potential Observations broadly describe the characteristics of student responses to tasks that should yield valid information about the students' attainment of the Focal KSAs. These Potential Observations are explicitly drawn from two assessment tasks presented in Table 2 of the Duncan, Rogat, and Yarden 2009 article (p. 668). The tasks exemplify how students may demonstrate progress toward sophisticated understanding of “the central role of proteins in genetic phenomena” (Duncan, Rogat, and Yarden, 2009, pp. 668-669). The tasks also appear in the *design pattern*, in the Exemplar Tasks field. The Potential Work Products are examples of what students may produce in response to assessment prompts. For the sake of illustration, there are in the *design pattern* two Potential Work Products for every Potential Observation in order to demonstrate how different types of work products can be prompted that can yield the same broad types of observations of student understanding. An observation is a quality of student work evidencing the Focal KSA; providing a range of Potential Work Products helps a task designer keep a focus on the substantive meaning of evidence, rather than the particular form. A given observation, therefore, may be evoked in multiple-choice format, open-ended responses, or as part of a larger investigation.

In the *design pattern*, there are in addition optional features that represent examples of varying attributes that could be designed into assessment tasks measuring attainment of the Focal KSAs. These *construct irrelevant features* can be varied because they are not about measuring student progress on the focal constructs of the learning progression, yet they accompany the construct-relevant features. The Additional KSA's: are examples of what additional knowledge skills and abilities may be required of students doing assessment tasks that are designed to measure their attainment of the Focal KSAs.

For the purpose of illustration, there is for each grade level band in the design band one Focal KSA, one Additional KSA, one Potential Observation, two Potential Work Products, and one scoring criterion per each of the two Exemplar Tasks. Most of the Variable Features, however, are relevant to more than one of the grade bands because they describe ranges of types of task stimuli that can be used to prompt more or less sophisticated levels of understanding about the focal concepts.

In the interactive version of the design pattern (http://design-drk.padi.sri.com/padi/AddNodeAction.do?NODE_ID=2233&state=viewNode), there are in addition Detail links to direct quotes from Duncan, Rogat, and Yarden (2009) that provide more background and context. Lastly, the References section contains citations to the main publication (Duncan, Rogat, & Yarden, 2009) plus references to related scholarly works cited in the Details.

Table 1. Illustrative Design Pattern Based on Genetics Learning Progression

Title	Illustrative design pattern based on genetics learning progression for grades 5-10
Overview	This design pattern describes students' evolving knowledge of the characteristics and functions of genes. Its contents are based on a published journal article by Duncan, Rogat, and Yarden (2009) in which the authors posit a learning progression for deepening students' understandings of modern genetics across grades 5-10. This understanding of modern genetics is identified in the paper as consisting of understanding of the genetic model, the molecular model, and the meiotic model. The paper posits 8 main ideas about the three models followed by the characteristics of what students in grade bands 5-6, 7-8, and 9-10 are capable of understanding about the main ideas respectively. **Details**
Use	Use this design pattern to build assessment arguments devoted to measuring the progression of student understanding about genetics in grades 5-10.
Focal KSAs	The Focal KSAs in this pattern correspond to increasing levels in the learning progression: 1. Understanding genes as "informational entities" (Duncan, Rogat, & Yarden, p. 665), present "in most cells in the organism" (Ibid., p. 664), that genes contain instructions for the growth and functioning of all living things, and that "Our body has multiple levels of organization, hence changes at one level may affect another" (Ibid, p. 660). (Grades 5-6) 2. Understanding that the genetic content specifies "very small biological entities" (proteins) "that carry out the functions in living things" (Ibid., p. 665), that proteins have "shapes and properties that afford their functions" (Ibid, p. 660), that changes to proteins can result from changes to genes and that those changes can "affect...structures and functions in the whole organism" (Ibid., p. 660) (Grades 7-8) **Details** 3. Understanding the "molecular processes involved in the translation of the genetic instructions into proteins" (Ibid., p. 666), understanding some of the" molecular structures of proteins (such as charge and size)" (Ibid., p. 666), and developing more sophisticated understandings of "genetic mutations...and their biological consequences at the molecular and cellular levels" (Ibid., p. 666) (Grades 9-10) **Details**
Additional KSAs	*Knowledge of different species of organisms that may be cited in the student tasks. (All bands)* *Knowledge of parts and functions of different organisms (Grades 5-6)* *Knowledge of different types of physiological functions that are genetically derived (Grades 7-8)* *Foundational knowledge about the structures and functions of molecules (Grades 9-10)*
Characteristic features of tasks	*Tasks prompt students to apply principles of cellular and/or molecular biology at grade-level appropriate levels of sophistication in order to give reasonable explanations or make reasonable predictions about the characteristics of genes, proteins, and the outcomes of genetic changes on cells and organisms. In other words, a task meant to determine whether a student is thinking at or above a specified level should present a situation to understand or explain such that the concepts described in the Focal KSAs for the level are required.* **Details**
Variable features of tasks	*Which types of organisms to focus on (Grades 5-10)* *Which types of cell structures to focus on (Grades 5-10)* *Whether to focus on normally varying traits such as eye color or different types of healthy vs. pathological genetic expressions (Grades 5-10)* *Which types of information representations to use , such as text, model diagrams, tables (Grades 5-10)* *Which types of mutations to focus on (Grades 7-10)* *Which types of proteins to focus on (Grades 7-10)* Which types of tissues, or organs to focus on (Grades 9-10)

Potential observations	Accuracy of information explicitly or implicitly provided in the student response about how: • "the alteration of a cell's structure or function can affect the structure or function of the organ or organism it resides in" (Ibid., p. 668) (Grades 5-6) • a change in a protein's shape might affect the protein's function, as well as the structure and function of a cell it resides in, and that of the whole organism" (Ibid., p. 668) (Grades 7-8) • a genetic mutation might influence the function or appearance of an organism by affecting the function or structure of a protein that acts within a cell, which resides in a tissue, and which functions in an organ" (Ibid., p. 668) (Grades 9-10)
Potential work products	*A list of inherited traits of different types of plants and animals (Grades 5-6)* *A narrative that differentiates between traits that result from alterations to cell structures and functions and those that result from mutations induced by infections or other environmental influences (Grades 5-6)* *A narrative describing examples of specific cellular changes that affect a body part or entire organism (Grades 7-8)* *Causal diagram of a model showing directionally appropriate cause and effect relationships between a particular type of genetic mutation and corresponding changes to cells and to the whole organism (Grades 7-8)* *Before and after sketches showing different types of change in different types of proteins and the impacts on of the changes on cell structures (Grades 7-8)* *Short narrative identifying a particular type of genetic mutation (Grades 9-10)* *Describing a plan to research at the molecular level the evolutionary relationships between two specific organisms (Grades 9-10)* *Report comparing and contrasting how doctors diagnose infectious diseases differently from genetically-inherited diseases in a way that reveals student understanding of the impacts of genetic mutations on cell structures (Grades 9-10)*
Exemplar tasks	Duncan, Rogat, and Yarden (2009) cite two assessment task examples: *Task description:* "Some people are born with a genetic disease called muscular dystrophy. People with this disease have great difficulty in walking or exercising. Can you explain what might be causing these problems? (Ibid., p. 668)? *Expected responses:* Grades 5-6: "Maybe these people have muscle cells that do not work well or maybe they have fewer muscle cells" (Ibid., p. 668). Grades 7-8: "Maybe their muscle cells do not move well because the proteins in these cells do not work well" (Ibid., p. 668). Grades 9-10: "Maybe their muscle cells do not move well because the proteins in these cells do not work as a result of a mutation in a gene" (Ibid., p. 668). *Task description:* "There is a protein called hemoglobin found in red blood cells that binds oxygen. It is possible that gene mutations could arise that prevents hemoglobin from binding oxygen. Explain how a mutation could cause this problem" (Ibid., p. 668). *Expected responses*: Grades 5-6: Not applicable. Grades 7-8: "Maybe a protein in the cell is changed so the cell cannot carry oxygen" (Ibid., p. 668). Grades 9-10: Maybe the hemoglobin protein is changed in shape, because of a mutation in a gene, so that hemoglobin cannot bind oxygen" (Ibid., p. 668).

References	Berenfeld, B., Damelin, D., Pallant, A., Tinker, B., Tinker, R., Xie, Q. (2004). *Molecular Workbench*. The Concord Consortium. Retrieved February 23, 2010. http:/ /www.concord.org. Duncan, R.G. (2006). Fostering generative understandings about complex phenomena in genetics. In: Barab S.A., Hay K.E., & Hickey D.T. (Eds.). *Proceedings of the Seventh International Conference for the Learning Sciences: Making a Difference*. Bloomington, Indiana (pp. 119–120). Mahwah, NJ: Erlbaum. Duncan, R.G., Rogat, A.D., Yarden, A. (2009) A Learning Progression for Deepening Students' Understandings of Modern Genetics Across the 5th–10th Grades. *Journal of Research in Science Teaching*, 46(6), 655-674. Duncan, R.G., & Reiser, B.J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching, 44(7),* 938–959. Duncan, R.G., Ruppert, J., Bausch, A., Freidenreich, H.B. (2008). *Promoting middle school students' understanding of molecular genetics.* Baltimore, MD: Paper presented at the Annual Meeting of the National Association for Research in Science Teaching. Krajcik, J., McNeill, K., & Reiser, B.J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education, 92(1),* 1–32. Rogat, A., Krajcik, J.S. (2006). *Supporting students understanding of current genetics in high school*. San Francisco: Paper presented at the Annual Meeting of the National Association for Research in Science Teaching. Roseman, J., Caldwell, A., Gogos, A., Kurth, L.A. (2006). *Mapping a coherent learning progression for the molecular basis of heredity.* San Francisco, CA: Paper presented at the Annual Meeting of the National Association of Research in Science Teaching. Venville, G., & Donovan, J. (2005). Searching for clarity to teach the complexity of the gene concept. *Teaching Science, 51(3),* 20–24. Venville, G., & Treagust, D.F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching, 35(9)*, 1031–1055.

Hence, through a process of content selection and inferential interpretation, the *design pattern* in Table 1 moves from elements of the learning progression into elements of a student reasoning and content domain, which can then support the development of valid assessment tasks and scoring procedures. These tasks and procedures can make evident (1) characteristics of student responses that are relevant for determining the students' progress on the learning progressions' assessable constructs and (2) characteristics of student responses that are irrelevant to the differentiation yet require prior knowledge or skill in order to be completed successfully. The *design pattern* can also serve to expose gaps in the learning progression that render it not yet sufficiently comprehensive for capturing the Student Model in enough comprehensiveness to be useful for designing assessment tasks and differentiating between their construct relevant and construct irrelevant characteristics.

2.2 Learning Progression Assessment Task Examples

Tables 2-4 show examples of tasks for each grade level band that illustrate how tasks can be designed through the attributes of the *design pattern* — and, hence, aligned to the learning progression. In task development, an evidentiary argument for the presence, level, or nature of targeted student knowledge expressed in a Focal KSA is constructed through logically connected claims or propositions, supported by data through warrants, and subject to alternative explanations. Warrants posit how responses in situations with the noted features depend on proficiency, which in this case is provided by the learning progression research. It is these situations that are prompted through assessment tasks. Each illustrative task in Tables 2–4 posits an assessment argument that suggests what students with certain levels of attainment of the knowledge expressed in the Focal KSA should be able to able to do successfully in various types of tasks. Low student performance on such tasks could also be due to inadequate additional skills or understandings that the student would also need for successful task completion that are not focal to the claim, such as knowledge of parts and functions of different organisms.

The process of constructing these assessment tasks involves iterating about how the Focal KSA for the targeted grade range implies a warrant for inferring a student’s level, based on what kinds of performances in what kinds of situations. This iteration moves around alternatives for task design in light of the warrant requirements expressed in the Focal KSAs, the Characteristic Feature of tasks and Potential Observations – the essential elements of the Messick quote. The resulting task for Grades 5-6 exhibited in Table 2 relies on student knowledge of parts and functions of the human body to respond correctly to a task designed to address the Focal KSA: "Understanding genes as informational entities present in most cells in the organism, that genes contain instructions for the growth and functioning of all living things, and that our body has multiple levels of organization, hence changes at one level may affect another." This knowledge of parts and functions of the human body instantiations demand for the Additional KSA: "Knowledge of parts and functions of different organisms." The task also represents the culmination of a decision to focus on the student comparing and contrasting healthy (i.e. blue eyes) and pathological (i.e. skin cancer) genetic expressions. This comparing and contrasting is expressed in the *design pattern* as a Variable Feature.

It was relatively easy for the task author to align Additional KSAs and Variable Features to the design of the three tasks because these attributes are inferable, implicit extensions of the Student

Model in the learning progression article. It was relatively more difficult however to identify characteristics of the intended observations in the student responses, because the Potential Observations expressed in the *design pattern* were derived directly from the learning progression paper. The task author addressed these constraints before moving on to consider Additional KSAs and Variable Features in the task design process, both of which are add-ons with reasonably inferable relationships to the explicitly-stated aspects culled from the learning progression paper. In the case of the Grades 5-6 task, the very specific attention to accuracy of information in the student response about how the alteration of a cell's structure or function can affect the structure or function of the organ or organism it resides in drove the construction of a prompt asking a student to compare how a person can end up with skin cancer to how a person can end up with blue eyes. The resultant observation that provides evidence of student possession of the Focal KSA would be a demonstration, through a successful comparison-contrast response, of his or her accurate knowledge about how gene expressions occur differently in meiosis compared to mitosis. Lastly, the task author needed to postulate grade-level appropriate work products for the tasks from the options expressed in the design pattern for the grade level range. This again was relatively easy because the Potential Work Products are also inferred *design pattern* attributes rather than construct-relevant constraints explicitly drawn from the learning progression article. Hence, for the grades 5-6 task, it was decided to have the task direct the student to compose a narrative that differentiates between a purely genetically-derived trait (blue eyes) that results from meiosis and a trait (skin cancer) that results from a genetic mutation and spreads through mitosis and that is at least partly the effect of the body's exposure to sunlight, an environmental condition.

The grades 7-8 task directs the student to "select one plant or animal and describe how it comes to be that all cells in an organism have the same DNA, yet perform different functions for the organism," then to "describe three examples of the DNA-derived functions for the organism you select." To provide a warrant for the Focal KSA " Understanding that the genetic content specifies very small biological entities (proteins) that carry out the functions in living things, that proteins have a shape (e.g., round, flat, spring-like) which affords their function, that changes to proteins can result from changes to genes, and that those changes can affect structures and functions in the whole organism," the task requires that the student possess the Additional KSA of knowing "the physiological features of an organism that are genetically derived." Yet, as identified in an aligned Variable Feature from the *design pattern* (i.e., "which types of organism to focus on"), the student is permitted in the task to select an organism on which to build his or her response. The scoring criteria are reflected by the Potential Observation ("Accuracy of information explicitly or implicitly provided in the student response about how a change in a protein's shape might affect the protein's function, as well as the structure and function of a cell it resides in, and that of the whole organism."), which is made possible through the scoring of the work product, "a narrative describing examples of specific cellular changes that affect a body part or entire organism."

The grades 9-10 task directs the student to explain "what happens at the molecular level that results in a particular type of mutation," and to "choose a mutation...to focus on." To provide a warrant for the Focal KSA "Developing more sophisticated understandings of genetic mutations and their biological consequences at the molecular and cellular levels," the task requires that the student possess the Additional KSA of "foundational knowledge about structures and functions of molecules." Yet, as identified in two aligned Variable Features from the *design pattern* (i.e., which types of cell structures and mutations to focus on), the student is permitted to select which type of mutation about which to build a response, knowing that different mutations impact different cell structures and functions, which in turn impact the functioning of the entire organism. The scoring criteria are reflected by the Potential Observation ("Accuracy of information about how a particular type of genetic mutation might influence the function or appearance of an organism by affecting the function or structure of a protein that acts within a cell, which resides in a tissue, and which functions in an organ"), which is made possible through the scoring of one of several Potential Work Products (diagram, narrative explanation, and slide presentation about what happens).

Table 2. Illustrative Task for Grades 5-6 Aligned to Design Pattern about Genetics Learning Progression

TASK
Prompt Compare how a person can end up with skin cancer to how a person can end up with blue eyes.
General attributes of a high-quality response Explains that skin cancer comes from a pathological form of genetic mutation from sunlight which, when untreated, passes to other cells in the body through mitosis. In contrast, blue eyes arise from the inheritance through meiosis of multiple genes that together code for the depositing in the iris of less pigment than is deposited to make brown eyes, hence allowing in the blue-eyed person more light to enter the eye than is possible when the person has brown eyes.
ALIGNMENT TO DESIGN PATTERN
Focal KSA Understanding genes as informational entities present in most cells in the organism, that genes contain instructions for the growth and functioning of all living things, and that our body has multiple levels of organization, hence changes at one level may affect another
Additional KSA Knowledge of parts and functions of different organisms
Characteristic feature of task Students apply principles of cellular and/or molecular biology at grade-level appropriate levels of sophistication in order to give reasonable explanations or make reasonable predictions about the characteristics of genes, proteins, and the outcomes of genetic changes on cells and organisms.
Variable feature of task Focus on different types of healthy vs. pathological genetic expressions
Potential observation Accuracy of information explicitly or implicitly provided in the student response about how: the alteration of a cell’s structure or function can affect the structure or function of the organ or organism it resides in
Potential work product A narrative that differentiates between traits that result from alterations to cell structures and functions and those that result from mutations induced by infections or other environmental influences

Table 3. Illustrative Task for Grades 7-8 Aligned to Design Pattern about Genetics Learning Progression

TASK
Prompt Select one plant or animal and describe how it comes to be that all cells in an organism have the same DNA, yet perform different functions for the organism. Describe three examples of such DNA-derived functions for the organism you select.
General attributes of a high-quality response Describes how DNA replicates throughout all cells in an organism through mitosis, yet different cells have distinctive functions for the organism and each function is ultimately the result of the synthesis of proteins, which is a process directed by the DNA and that many of the proteins, acting as enzymes, catalyze chemical reactions that result in the specialized cellular functions.
ALIGNMENT TO DESIGN PATTERN
Focal KSA Understanding that the genetic content specifies very small biological entities (proteins) that carry out the functions in living things, that proteins have a shape (e.g., round, flat, spring-like) which affords their function, that changes to proteins can result from changes to genes and that those changes can affect structures and functions in the whole organism.
Additional KSA Knowledge of different types of physiological functions that are genetically derived
Characteristic feature of task Students apply principles of cellular and/or molecular biology at grade-level appropriate levels of sophistication in order to give reasonable explanations or make reasonable predictions about the characteristics of genes, proteins, and the outcomes of genetic changes on cells and organisms.
Variable features of task Which types of organisms to focus on Which types of cell structures to focus on Which types of information representations to use, such as text, diagrams, tables
Potential observation Accuracy of information explicitly or implicitly provided in the student response about how a change in a protein's shape might affect the protein' s function, as well as the structure and function of a cell it resides in, and that of the whole organism.
Potential work product A narrative describing examples of specific cellular functions that affect a body part or entire organism

Table 4. Illustrative Task for Grades 9-10 Aligned to Design Pattern about Genetics Learning Progression

TASK
Prompt What happens at the molecular level that results in a particular type of mutation? Choose a mutation that you want to focus on.
General attributes of a high-quality response Correctly diagrams and explains a specific type of error in the replication of nucleotides in a cell's DNA molecule during the synthesis phase of cell division.
ALIGNMENT TO DESIGN PATTERN
Focal KSA Developing more sophisticated understandings of genetic mutations and their biological consequences at the molecular and cellular levels
Additional KSA Foundational knowledge about the structures and functions of molecules
Characteristic feature of tasks Students apply principles of cellular and/or molecular biology at grade-level appropriate levels of sophistication in order to give reasonable explanations or make reasonable predictions about the characteristics of genes, proteins, and the outcomes of genetic changes on cells and organisms.
Variable feature of tasks Which types of cell structures to focus on Which types of mutations to focus on
Potential observation A particular type of genetic mutation might influence the function or appearance of an organism by affecting the function or structure of a protein that acts within a cell, which resides in a tissue, and which functions in an organ
Potential work product Diagram of what happens Narrative explanation of what happens Slide presentation about what happens

2.3 Using ECD to Demarcate Learning Progression Construct Boundaries

Use of ECD with learning progression assessment design requires identifying how the constructs of the learning progression are to be demarcated with regard to the progression's internal subcomponents and to its relationships with other progressions. Different learning progression research projects vary in how far along they are in making explicit their demarcations. Decision-making about demarcation is complicated by the possibility that a subcomponent of one learning progression may be its own distinct learning progression. Corcoran, Mosher, and Rogat (2009) maintain that "It... is reasonable to think of any particular progression as being made up of sets of component progressions, each of which could be specified in a similar way" (p. 42). Applying ECD, one must be able to identify the bounded interdependent characteristics of student thinking and doing that deserve to be captured in one learning progression-based *design pattern*, yet also be prepared to identify cases where characteristics are better modeled in related, yet separate, *design patterns*.

In cases of the latter, it is important to identify what relationships connect the components of the learning progressions represented in the *design patterns*. For example, a *design pattern* capturing the characteristics of a learning progression about scientific reasoning may be connected to a *design pattern* capturing the characteristics of a learning progression about reading or writing or mathematical reasoning; or, a *design pattern* about a certain component of a learning progression about domain specific scientific reasoning may be connected to other *design patterns* about other components of scientific reasoning in that same learning progression. These confluences of learning progressions and their components can be made evident through the structure of a *design pattern*. Another demarcation challenge for use of ECD with learning progressions is how to use the *design pattern* structure to support valid differentiation between science content constructs and inquiry skill constructs in the progression.

The deliberative *domain modeling* that is initiated through *design pattern* development can support learning progression researchers in their efforts to build assessments that can be used to test alternative representations of the components of learning progressions and their interdependent relationships. For example, through *design pattern* development, there can be iterative phases of *domain modeling* that explore treating content understandings and inquiry practices in the learning progression separately, then together. *Design patterns* lend themselves well to such iterative work because a *design pattern* is structurally agnostic on what constitutes appropriate demarcation of learning progression components, yet is capable of capturing, for example, one level of a learning progression, one construct across levels, one construct per level, or a whole network of interdependent learning progressions. The *design pattern* in Table 1, for example, took the most comprehensively-articulated construct provided in a published article to render inferential judgments that support assessment argument-building on that construct. Decisions about how granular other particular learning progression-based *design patterns* should be would depend not only on what characterizes the boundaries of assessable constructs within the progression, but also on how much is known about how the learning of the construct evolves. The process of differentiating between Focal and Additional KSAs would help. Focal forms of knowledge, skills, and abilities (KSAs) can express the anticipated levels of development of student reasoning on the targeted learning progression and Additional KSAs can identify additional characteristics of student knowledge or skill from different intersecting learning progressions. An accompanying benefit could be that *design pattern* development could be informative of whether the additional knowledge and skill elements should be the foundation of other distinctive Student Models and, hence, deserving of their own *design patterns*.

2.4 The Usefulness of Evidence–Centered Design for Aligning Learning Progressions and Science Standards

If we are to assume that certain types of instructional programs are more amenable to fostering student growth on a learning progression than others, it would be appropriate to revise accountability standards to more accurately reflect and support instructional redesigns. As expressed by Corcoran, Mosher, and Rogat (2009), "states and districts revising their standards and trying to improve science teaching would benefit from considering the lessons (the learning progressions) provide about the sequencing of the science curriculum, the interconnections between conceptual understanding and practices, and the design of assessments" (p. 52).

This redesign process would need to take account of how state accountability standards and benchmarks often do a better job of articulating the characteristics of intended student behaviors than they do the characteristics of student reasoning. Typically, but not always, the desired outcomes are expressed as student abilities to carry out certain activities and express certain facts or concepts, and less often do they specify criteria for judging intermediate levels of understanding or quality of explanation. Though it is fair to say that standards do not undercut an instructor who wants to assign a high value to a student response that demonstrates incomplete yet evolving levels of reasoning, it is also fair to say that too often the benchmarks, with their focuses on behavior, are wide enough to permit the utilization of assessment prompts and stimuli that can elicit the behaviors called out in the benchmarks without requiring the exhibitions of reasoning that would permit identification of where the student is on the learning progression.

Table 1 illustrated how ECD is capable of bridging learning progressions and standards because, through *design patterns*, intermediate levels in the evolution of skills and understandings (which culminate in deep, full-fledged canonical reasoning capabilities at the upper levels of the learning progressions) can be modeled and hence identified for the design of instructional and assessment tasks that recognize and assign a value to evolving progressions of reasoning and understanding. In addition, ECD can aid in the redesign of standards and benchmarks and their subsequent greater alignment to learning progressions and learning progression-focused instructional programs by supporting greater differentiation among the characteristics of student reasoning at different grade levels.

The first step, however, is to discern what alignments already exist. For the sake of illustration of the current state of alignment between a learning progression and a set of state standards, Table 5 contrasts some Minnesota science benchmarks on the topic of the movement of water through the environment with descriptions of evolving understandings from a learning progression article about the same topic (Gunckel, Covitt, & Anderson, 2009). The purpose of this exercise is not to judge the worth of one against the other but merely to show how differently they express student outcomes. The student outcomes in the learning progression are expressed as cognitive research-driven descriptions of student reasoning at various developmental levels, whereas the benchmarks express intended student behaviors. The key contrasting verbiage is italicized to show how the student behaviors are far more connected to reasoning characteristics in the learning progression than in the state benchmarks.

As Table 5 illustrates, as long as the benchmarks use process-oriented verbiage to describe student tasks without connecting the tasks to a stimulus, it is difficult to interpret results of student responses to the task in the context of the learning progression. For example, to cite the last benchmark in the bottom row of Table 5, if an assessment task asked a student to "*explain* how the rearrangement of atoms and molecules in a chemical reaction illustrates conservation of mass," the explanation could be the product of the student's memorization of information and not be helpful in differentiating level of reasoning. As such, it would hence not yield evidence about whether the student can reason at the level specified in the learning progression. To take another example, a task that calls on students to identify certain properties of materials may also either require recalling information or making simple inferences. If instead the task were designed to require use of model-based reasoning about evidence presented in the stimulus in order to answer correctly, then the task would succeed as a measure of a higher level of attainment on the learning progression.

ECD has an opportunity here to play a bridging role because, in a *design pattern*, Characteristic Features of tasks can be identified that capture the characteristics of the intended student thinking, and these Characteristic Features can be merged with Variable Features to design valid tasks in which the student behavior captured by the response provides a warrant for ascertaining the students' level of reasoning. The attributes of *design patterns* pertaining to products and observable behaviors can help to identify the requirements that the assessment tasks should fulfill if they are to yield evidence in the student behaviors about which level of reasoning students are capable of demonstrating at a particular point in time in their progression of learning about the focal construct.

Table 5. Illustrative Alignment between Learning Progression and Benchmarks

Learning progression descriptions (Gunckel, Covitt, and Anderson, 2009, pp. 11-12).	**Aligned Minnesota benchmarks**
At Level 2, Students still explain and predict using force dynamic reasoning, but are now giving more attention to hidden mechanisms in their accounts. They *recognize* that events have causes and often describe simple mechanisms that they use to explain or predict events. They are beginning to trace water and substances, *recognizing* that water and substances that are no longer visible go someplace else." Students still think about water as part of the background landscape, but their *conception* of the size of the background landscape is larger. Level 2 students *think about* rivers as connected to other rivers and groundwater as layers of water underground. Level 2 students *think about* the movement of water as a natural tendency of water and they identify possible enablers and antagonists to movement.)	The student will *observe* that water can be a solid or liquid and can change from one state to another (Structure of matter - Grade 2) The student will *identify* where water exists on earth (Structure of matter - Grade 4) The student will *describe* the water cycle involving the processes of evaporation, condensation, precipitation and collection (Structure of matter - Grade 4)
At Level 3, students are *recognizing* that water and substances in water are parts of connected systems and they can tell stories that use processes to move water and substances through systems. However, there *are gaps in students' reasoning* that suggests that students' stories are not connected into complete models that they use to explain and predict. This level represents *the beginning of model-based reasoning*.	The student will *define* chemical and physical changes (Chemical Reactions - Grade 6) The student will *give examples and classify* substances as mixtures or pure substances (Chemical Reactions - Grade 6)
At Level 3, students trace water through multiple pathways in connected systems. However, *the nature of the connections among systems is not always clear* to students)	The student will *identify* the forces that create currents and layers in the Earth's atmosphere and water systems (The Water Cycle, Weather in Climate - Grade 8) The student will *trace* the cyclical movement of carbon and water through the lithosphere, hydrosphere, atmosphere and biosphere (The Water Cycle, Weather in Climate - Grades 9-12) The student will *identify, predict and investigate* the factors that influence the quality of water and how it can be reused, recycled and conserved (The Water Cycle, Weather in Climate - Grades 9-12)
At Level 4, students *use scientific model-based accounts* to explain and predict. Their explanations connect observations to patterns and models and use appropriate models and principles. Their predictions use data about particular situations along with principles to determine the movements of water and substances in water. *Students who use scientific model-based thinking* can trace water and substances in water along multiple pathways through connected systems and describe these pathways and movements at multiple scales.	The student will *observe* that substances react chemically with other substances to form new substances with different characteristic properties (Chemical Reactions - Grade 6) The student will *describe* chemical reactions using words and symbolic equations (Chemical Reactions - Grade 6) The student will *explain* the influence of temperature, surface area, agitation and catalysts on the rate of a reaction (Chemical Reactions - Grades 9-12) The student will *explain* how the rearrangement of atoms and molecules in a chemical reaction illustrates conservation of mass (Chemical Reactions - Grades 9-12)

3.0. Applying ECD to the Validation of Learning Progressions

The same ECD-grounded assessments that can be used to measure student advancement through a learning progression-grounded instructional program can also be used to test the validities of the learning progressions themselves. This is because ECD provides a process for making careful alignments between assessment tasks and the models of student reasoning that are hypothesized in the learning progressions. Then, assuming resultant face validity between the assessment tasks and the learning progression constructs, and assuming sufficient student opportunity to learn, psychometric analyses of student results can be used to test the legitimacy of the Student Models postulated in the learning progression in more valid ways than would be possible if non-ECD-based assessment results were used instead.

There are several dimensions to the challenge of validating learning progressions. First, capabilities for validating learning progressions are undermined by lack of consensus about (1) to what extent domain specific learning progresses in a predictable course that corresponds to naturally evolving cognitive capacity and (2) to what extent different instructional practices help or hinder the progression. If a learning progression is responsive to generalizable evolutions of student cognitive capacity, it follows then that the underlining learning progression should be achievable among all students as long as their cognitive capacities are not hindered by poor instruction or other negative environmental influences. Yet, to what extent must the instructional program teach explicitly to the progression to ensure that students make progress — or, in other words, to what extent may a student progress irrespectively of the design of the instructional program? An analogy would be to physical growth. Children grow naturally as they get older, yet when a child is deprived of proper nutrition, this deprivation may contribute to stunted growth. It is clear that poor nutrition, just like poor instruction, can inhibit growth. Yet, accepting the premise that there is a natural evolution of student reasoning, the question of how circumscribed must a nutritional program be for a child to physically grow as large as he is capable of growing can also be asked about the construct learning addressed by the learning progression. Specifically, how circumscribed must an instructional program be for a child to progress as far as he is capable of progressing?

This lack of consensus about how learning progresses relative to schooling makes it more difficult to determine what empirical data need to be gathered to identify the characteristics of the learning progression. Can the characteristics be ascertained in some pure state independently of the mediating influences of teachers, parents, or other learning agents? Some learning progression researchers attempt to avoid these influences by asking cognitive psychologists and domain experts to delineate progression characteristics. Others, however, study student thinking and behavior in instructional settings and then use their findings to design instructional strategies that support the progression. For example, a team of researchers who developed a multiyear curriculum that teaches to a particular middle school-level learning progression known as *Investigating and Questioning our World through Science and Technology* designed assessments to administer to students as they progress through the years of the curriculum so that they could study how the students' understandings evolved over time. Data from these assessments informed revising and refining the progression (Krajcik, McNeill, & Reiser, 2008). Another team of researchers carried out a cross-sectional rather than longitudinal investigation. They conducted clinical interviews of different groups of students at different grade levels in order to gather evidence for how learning about the focal content was progressing among separate grade-level cohorts of students taking different courses with different teachers (Mohan, Chen, and Anderson, 2009). Yet, it is fair to surmise that because instructional practices in different classrooms settings will not be consistent, there is risk that the characteristics of instruction in the students' current and prior learning experiences may confound these researchers' attempts. As noted by Duncan and Hmelo-Silver (2009), "a valid progression implies that the underlying cognitive model of learning holds true in different instructional settings and for different learners. However, learners bring with them unique experiences and knowledge and it is not yet clear how learning progressions can take into account these different learners histories" (p. 608).

Further challenging our ability to delineate valid learning progressions is conflicting evidence from psychometric analyses of responses to assessments that are designed to identify where students fit on a progression. Some researchers report confirmatory results. For example psychometric analyses (e.g., factor analyses, cross-sectional analyses, growth curve analyses) of responses to constructed response items designed to yield evidence of student level on the biodiversity learning progression supported the assessment's construct validities. Yet, assessment results on items designed to validate a learning progression about force and motion did not upon analysis appear to provide such validation. Latent class analyses of responses to "ordered multiple choice" items (Briggs et al., 2006) were performed. The items contained selection choices designed to correspond to different learning progression levels. Researchers did not find consistent student performances across the items that should have been correlated if the learning progression was valid (Steedle & Shavelson, 2009). It was not clear from the findings if the lower than expected correlations were due to inadequate item construction or to structural problems with the learning progression.

Hence, learning progression research and development is challenged by the need for greater clarity about (1) what aspects of knowledge and skills characterize the progressions, (2) what characterizes their relationships to natural cognitive development in contrast to instructional programs and other environmental influences, (3) what characterizes the nature of what instructors can do to help or hinder student progress, and (4) how the dimensions of cognitive capabilities expressed in the progressions can be more clearly mapped to the characteristics of knowledge and forms of inquiry held canonically by the different scientific communities. To answer these fundamental questions, validations are needed of the methods used by learning progression researchers to gather the evidence to demarcate their progressions' boundaries and levels.

4.0 Conclusion

Utilization of ECD and *domain modeling* through *design pattern* construction can support the design of assessments that can measure both student progress in learning progression-based instructional programs and also support learning progression validation efforts. Then, assuming that increasing numbers of instructional programs get redesigned in constructivist directions that respond to learning progressions by rewarding deep reasoning more than correct answers, ECD can be used to revisit and, if need be, revise accountability standards to more directly support and reward this pedagogical paradigm shift. ECD is well suited to these efforts because it provides a systematic, deliberative process for designing valid assessment tasks. Yet, being that the tasks are rooted in Student Models and being that the models are only as valid as the assessment evidence that supports them, these utilizations of ECD will only be useful for assessing student progress if the progression itself truly captures how student reasoning evolves into substantive acquisition of canonical knowledge and skills.

References

American Association for the Advancement of Science. (2006). *AAAS Project 2061 Biology Textbooks Evaluation*. Retrieved January 25, 2010. http://www.project2061.org/publications/textbook/hsbio/summary/genome.htm.

Briggs, D.C., Alonzo, A.C., Schwab, C., Wilson, M. (2006). Developmental assessment with ordered multiple-choice items. *Educational Assessment, 11(1),* 33–63.

Corcoran, T., Mosher, F.A., Rogat, A. (2009). *Learning progressions in science: An evidence-based approach to reform.* Teachers College-Columbia University: Consortium for Policy Research in Education.

Duncan, R.G., Hmelo-Silver, C.E., (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching, 46(6),* 606-609.

Duncan, R.G., Rogat, A.D., Yarden, A. (2009). A Learning Progression for Deepening Students' Understandings of Modern Genetics Across the 5th–10th Grades. *Journal of Research in Science Teaching, 46(6),* 655-674.

Gunckel, K.L., Covitt, B.A., Anderson, C.W. (2009) Learning a secondary discourse: shifts from force-dynamic to model-based reasoning in understanding water in socio-ecological systems. *Paper presented at the Learning Progressions in Science (LeaPS) Conference, June 2009, Iowa City, IA*

Kennedy, C.A., Wilson, M. (2007). Using progress variables to interpret student achievement and progress. *BEAR Report Series, 2006-12-01.*University of California, Berkeley.

Krajcik, J., McNeill, K. L., Reiser, B. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate Project-Based Pedagogy*. Science Education, 92(1)*, 1-32.

Mislevy, R., Hamel, L., Fried, R., G., Gaffney, T., Haertel, G., Hafter, A., Murphy, R., Quellmalz, E., Rosenquist, A., Schank, P., Draney, K., Kennedy, C., Long, K., Wilson, M., Chudowsky, N., Morrison, A., Pena, P., Songer, N., Wenk, A. (2003). *Design patterns for assessing science inquiry (PADI Technical Report 1).* Menlo Park, CA: SRI International.

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice*, *25*(4), 6–20.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1,* 3-67.

Mohan, L., Anderson, C.W. (2009). Teaching experiments and the carbon cycle learning progression. *Paper presented at the Learning Progressions in Science (LeaPS) Conference, June 2009, Iowa City, IA.*

Mohan, L., Chen, J., Anderson, C.W. (2009). Developing a K-12 learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46(6), 675-698.*

Schwarz, C.V., Reiser, B.J., Davis, E.A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46(6), 632-654.*

Smith, C., Wiser, M., Anderson, C.W, Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic molecular theory. *Measurement: Interdisciplinary Research and Perspectives, 14(1&2)*, 1-98.

Songer, N.B., Kelcey, B., Gotwals, A.W. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about diversity. *Journal of Research in Science Teaching, 46(6), 610-631.*

Steedle, J.T., Shavelson, R.J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching, 46(6), 699-715.*

!

Sponsor
The National Science Foundation, Grant DRL - 0733172

Prime Grantee
SRI International. *Center for Technology in Learning*

Subgrantees
University of Maryland
Pearson